Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	1994	Journal article

4. TITLE AND SUBTITLE

Identification and characterization of a gene cluster mediating enteroaggregative Escherichia coli aggregative adherence fimbria I biogenesis

5. FUNDING NUMBERS

PE - 61102A
PR - 001.01
TA - BIX
WU - 1440

6. AUTHOR(S)

Savarino SJ, Fox P, Yikang D, Nataro JP

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Medical Research Institute
Commanding Officer
8901 Wisconsin Avenue
Bethesda, Maryland 20889-5607

8. PERFORMING ORGANIZATION REPORT NUMBER

NMRI 94-25

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Naval Medical Research and Development Command
National Naval Medical Center
Building 1, Tower 12
8901 Wisconsin Avenue
Bethesda, Maryland 20889-5606

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

DN244540

11. SUPPLEMENTARY NOTES

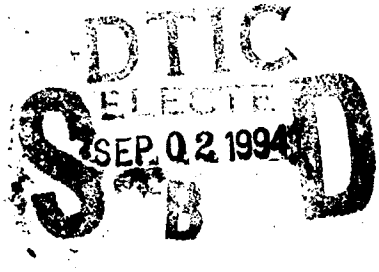
Reprinted from: Journal of Bacteriology 1994 Aug;Vol.176 No.16 pp.4949-4957

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*



14. SUBJECT TERMS

enterically transmitted diseases; enteroaggregative Escherichia coli; bacterial amino acid sequences; nucleotide sequences; fimbria

15. NUMBER OF PAGES

9

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

JOURNAL OF BACTERIOLOGY, Aug. 1994, p. 4949–4957
0021-9193/94/$04.00+0

Vol. 176, No. 16

Identification and Characterization of a Gene Cluster Mediating Enteroaggregative *Escherichia coli* Aggregative Adherence Fimbria I Biogenesis

STEPHEN J. SAVARINO,[1]* PATRICIA FOX,[1,2] DENG YIKANG,[3] AND JAMES P. NATARO[3]

Enteric Diseases Program, Naval Medical Research Institute, Bethesda,[1] GeoCenters, Inc., Fort Washington,[2] and Center for Vaccine Development, University of Maryland School of Medicine, Baltimore,[3] Maryland

Received 31 December 1993/Accepted 9 June 1994

The aggregative pattern of adherence (AA) exhibited by enteroaggregative *Escherichia coli* upon HEp-2 cells is a plasmid-associated property which correlates with aggregative adherence fimbria I (AAF/I) expression and human erythrocyte hemagglutination. By using cloning and mutagenesis strategies, two noncontiguous plasmid segments (designated regions 1 and 2) required for AA expression have previously been identified in enteroaggregative *E. coli* 17-2. Tn*phoA* mutagenesis was performed on clones containing region 1, and 16 Tn*phoA* mutants which were negative for the AA phenotype were analyzed. The Tn*phoA* insertion site for each mutant was determined by junctional DNA sequencing. All 16 mutations occurred within a 4.6-kb span in region 1. Nucleotide sequence analysis of the region revealed four contiguous open reading frames, designated *aggDCBA*, in the same span. AA-negative Tn*phoA* insertions into all open reading frames except *aggB* were obtained. On the basis of mutational analysis and protein homology data, it is inferred that *aggA*, *aggC*, and *aggD* are involved in biogenesis of AAF/I, encoding a major fimbrial subunit, outer membrane usher, and periplasmic fimbrial chaperone, respectively. By immunogold electron microscopy, polyclonal antiserum raised against the *aggA* gene product decorated AAF/I fimbriae, affirming that AggA encodes an AAF/I subunit.

Enteroaggregative *Escherichia coli* (EAggEC) strains have been incriminated as causative agents of persistent childhood diarrhea (9, 38). EAggEC strains are defined by their pattern of mannose-resistant adherence to HEp-2 or HeLa cells. Bacterial aggregates adhere to epithelial cells in a stacked-brick-like lattice, so-called aggregative adherence (AA) (30). The AA phenotype is distinct from the localized adherence (LA) and the diffuse adherence (DA) patterns of epithelial cell adherence characteristic of enteropathogenic *E. coli* (EPEC) and diffusely adherent *E. coli* (DAEC), respectively. The LA pattern is typified by the formation of distinct bacterial microcolonies or clusters adherent to HEp-2 cells, while for the DA pattern single bacteria adhere over the epithelial cell surface in an evenly dispersed arrangement. EAggEC strains adhere to human-derived intestinal explants and cultured intestinal mucosa in a pattern similar to that seen in the HEp-2 cell adherence assay (20, 47, 48), suggesting that this adherence phenotype may have relevance in human colonization and disease.

For several EAggEC strains, including prototype strain 17-2, the genetic determinants conferring the AA phenotype are associated with large, ca. 60-MDa plasmids, which show a high degree of conservation among EAggEC strains (29, 43). In EAggEC strain 17-2, two noncontiguous regions on its single large plasmid, termed regions 1 and 2, are required for expression of the AA phenotype, aggregative adherence fimbria I (AAF/I) expression, and mannose-resistant hemagglutination (31). Region 1 determinants required for the AA phenotype are contained on a 6.8-kb *Cla*I fragment (32) (see Fig. 1). A plasmid construct containing this fragment can be complemented in *trans* to express AAF/I fimbriae and the AA phenotype by a 1.2-kb fragment which encompasses region 2 (32). Mapping data show that regions 1 and 2 are separated by 9 kb on the plasmid of strain 17-2 (31), and the structural gene for a heat-stable enterotoxin is found in the interceding span (39).

In this paper, we present detailed mutational analysis of the AAF/I determinants in region 1 and the corresponding nucleotide sequence. The described gene cluster and its deduced protein products show organizational and structural similarities to other known fimbrial biogenesis systems of pathogenic gram-negative bacteria.

MATERIALS AND METHODS

Bacterial strains, plasmids, and growth conditions. EAggEC strain 17-2 (serotype O3:H2) is a Chilean pediatric diarrheal isolate which harbors a 60-MDa plasmid, p17-2, encoding AA determinants (43). All of the plasmid constructions used in this study are derived from p17-2 and are represented in Fig. 1A. To construct plasmid pSS101, Tn*A*, encoding ampicillin resistance, was introduced into p17-2 as previously described (35). *E. coli* HB101 was used as host for plasmid constructs. 17-2(Δ*aggR*) has an in-frame deletion in the *aggR* gene and does not express AAF/I (32).

Strains were routinely grown on LB medium (36) at 37°C and stored at −70°C in LB broth supplemented with 20% glycerol. Supplements were used in the following concentrations: streptomycin, 100 μg/ml; ampicillin, 200 μg/ml; kanamycin, 50 μg/ml; tetracycline, 30 μg/ml; and the alkaline phosphatase chromogenic substrate 5-bromo-4-chloro-3-indolyl-phosphate (XP) (Sigma), 40 μg/ml.

Recombinant DNA techniques. Plasmid DNA was purified by the method of Birnboim and Doly (6) or by cesium chloride-ethidium bromide density gradient centrifugation following alkaline extraction (36). Restriction endonucleases and other DNA-modifying enzymes were purchased from Boehr-

* Corresponding author. Mailing address: Enteric Diseases Program, Naval Medical Research Institute Annex, 12,300 Washington Ave., Rockville, MD 20852. Phone: (301) 295-1636. Fax: (301) 295-6171.

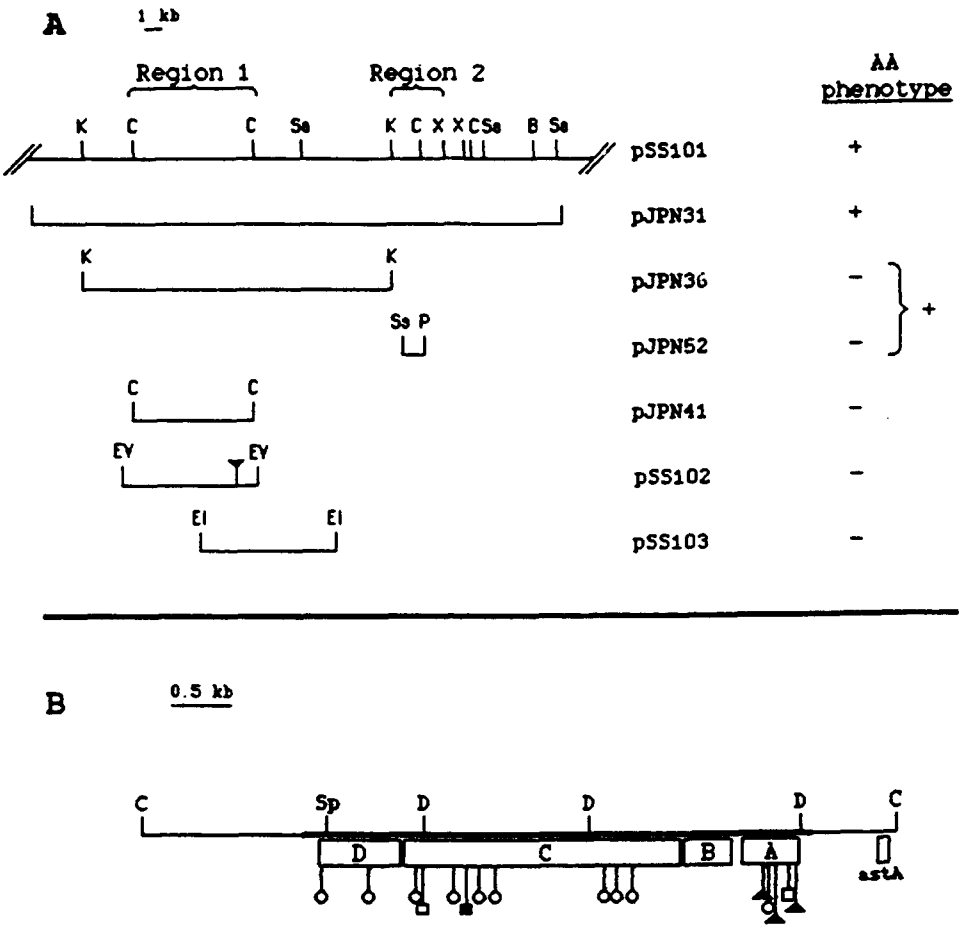


FIG. 1. (A) Restriction map of pSS101 showing the region harboring AAF/I regions 1 and 2 (demarcated by brackets). Relevant subclones are depicted below this map, with their corresponding adherence phenotypes indicated to the right. The cloning vectors (with the relevant construction in parentheses) were as follows: pCVD301 (pJPN31), pRK415 (pJPN36), pBR328 (pJPN52), pBluescript/SKI (pJPN41), and pUC19 (pSS102 and pSS103). The bracket grouping pJPN36 and pJPN52 indicates that these clones complement one another in *trans* to express the AA phenotype. (B) Enlargement of the 6.8-kb *Cla*I fragment representing region 1 and a physical map of the *agg* gene cluster. The positions of the individual genes are indicated by the lettered open boxes. Location of the EAST1 enterotoxin gene (*astA*) is also shown (39). Tn*phoA* insertion sites into region 1-containing clones which abolish the AA phenotype are marked by lollipops. Triangles indicate insertions into pSS101, squares indicate insertions into pJPN31, and circles indicate insertions into pJPN36. Solid lollipops signify PhoA$^+$ mutations, while open lollipops indicate PhoA$^-$ mutations. The nucleotide sequence for the 4.6-kb segment denoted by the thick black line is presented in Fig. 2. Restriction sites shown are *Cla*I (C), *Sph*I (Sp), *Dra*I (D), *Kpn*I (K), *Sal*I (Sa), *Xba*I (X), *Bam*HI (B), *Ssp*I (Ss), *Pvu*I (P), *Eco*RV (EV), and *Eco*RI (EI).

inger-Mannheim (Indianapolis, Ind.) and used according to the manufacturer's instructions. Conjugations were performed by the plate cross-streak method (36), and transformations were performed by the technique described by Mandel and Higa (27).

Tn*phoA* mutagenesis and determination of Tn*phoA* insertion sites. Tn*phoA* mutagenesis of the AA-conferring recombinant cosmid pJPN31 was previously reported (29), as was the assignment of these mutations to either region 1 or region 2 of AAF/I by restriction mapping (31). Additional transposon mutagenesis was performed separately on HB101(pSS101) and HB101(pJPN36/pJPN52), both of which express the AA phenotype (Fig. 1A). Tn*phoA* was introduced into these constructs by using the conjugative suicide vector pRT733 as previously described (42). For HB101(pSS101), each conjugation mixture was plated on agar supplemented with streptomycin, kanamycin, and XP. Blue colonies, indicating an in-frame Tn*phoA* insertion into a gene whose product is transported extracytoplasmically such that extracellular alkaline phosphatase (PhoA) activity is detectable colorimetrically, were pooled from each conjugation. To eliminate those mutants with chromosomal insertions, plasmid DNA was prepared from each pool and used to transform HB101. Transformants expressing PhoA activity on agar supplemented with kanamycin, streptomycin, and XP were studied for adherence capability in the HEp-2 cell adherence assay.

Tn*phoA* was similarly introduced into HB101(pJPN36). Transpositions into pJPN36 were detected by transforming pooled plasmid DNA from individual matings and selecting kanamycin-resistant (Km^r) transformants without initial regard for PhoA activity. pJPN36::Tn*phoA* mutants were screened for loss of the AA phenotype in the HEp-2 cell adherence assay in an HB101(pJPN52) host. All AA$^-$ mutants were subsequently tested for PhoA activity on XP-supplemented media.

DNA sequence analysis. The DNA sequence of region 1 was determined by the dideoxynucleotide chain termination method of Sanger et al. (37) on double-stranded templates by

using *Taq*I cycle sequencing kits (Applied Biosystems [ABI], Foster City, Calif.) in a thermal cycler (Perkin-Elmer Cetus). For sequence analysis, DNA was purified by the alkaline lysis-polyethylene glycol purification method recommended by the manufacturer of the sequencing kits (2). Templates included pSS102, pSS103, pJPN41 (Fig. 1A), and DNA fragments subcloned from these constructions in pUC18 or pUC19. Both strands of DNA were sequenced by using fluoresceinated forward and reverse pUC primers (ABI) with primer chemistries or custom oligonucleotide primers with terminator chemistries. Primers were synthesized as needed on an ABI model 392 automated DNA/RNA synthesizer.

For each AA$^-$ transposon mutant, the Tn*phoA* insertion point was determined by DNA sequence analysis. An *Eco*RV or *Sal*I DNA fragment containing all or part of the transposon, respectively, and flanking DNA were subcloned into pUC18 or pBluescript/SKI by using kanamycin selection. To sequence outward from Tn*phoA*-containing inserts, a primer complementary to the 5′ end of the truncated *phoA* gene (5′-TCGCTAAGAGAATCA...3′) was used.

The MacVector Sequence Analysis software program (International Biotechnologies, New Haven, Conn.) was used to determine restriction enzyme sites, open reading frames (ORFs), and deduced amino acid sequence from the obtained DNA sequence. Protein database searches, comparisons, and alignments were performed with MacVector software by using the GenBank Entrez Protein Database.

Construction of a *malE-aggA* fusion and generation of anti-AggA polyclonal antiserum. The *aggA* gene lacking its predicted signal sequence was amplified with the oligonucleotides 5′-GCGTTAGAAAGACCTCCAATA-3′ (sjs5) and 5′-GCCGGATCCTTAAAAATTAATTCCGGC-3′ (sjs6) as primers (see Fig. 2) and plasmid pSS103 (Fig. 1A) as the template with the GeneAmp PCR kit (Perkin-Elmer Cetus). The PCR cycle included denaturation for 1 min at 94°C, primer annealing for 1 min at 30°C, and extension for 1 min at 72°C (30 cycles). The amplification product was digested with Klenow fragment and *Bam*HI and cloned into the expression vector pMAL-p2 (New England Biolabs [NEB], Beverly, Mass.), which had been digested with *Xmn*I-*Bam*HI. In this way, the truncated *aggA* gene was inserted downstream from the *malE* gene with its signal sequence. This construction, designated pSS120, was confirmed by DNA sequence analysis with the *malE* primer (NEB) and transformed into *E. coli* TB1 for expression and purification of maltose-binding protein (MBP)-AggA fusion protein (3).

Cultures of TB1(pSS120) were grown in rich medium with glucose (3) at 37°C with shaking to an optical density at 600 nm of 0.5, followed by induction with 0.3 mM isopropyl-β-D-thiogalactoside (IPTG) (Sigma) for 2 h. Osmotic-shock preparations were made from washed cells, and the periplasmically expressed MBP-AggA fusion was purified by affinity chromatography on a cross-linked amylose resin column, following the manufacturer's (NEB) recommendations.

New Zealand White male rabbits (1.5 kg) were immunized intramuscularly with purified MBP-AggA fusion mixed 1:1 with complete Freund's adjuvant. At 2- to 4-week intervals, the rabbits received four additional doses of MBP-AggA mixed 1:1 in incomplete Freund's adjuvant. Three weeks after the last immunization, the animals were phlebotomized to obtain serum.

HEp-2 cell adherence assay. The HEp-2 cell adherence assay was performed according to the original method of Cravioto et al. (8). Briefly, 20 μl of an overnight bacterial culture (grown statically in LB broth at 37°C) was added to the HEp-2 cell monolayer (50 to 75% confluence) in 24-well tissue culture plates (Costar Corporation, Cambridge, Mass.). Infected monolayers were incubated in Minimal Essential Medium (Life Technologies, Grand Island, N.Y.) containing 0.5% D-mannose for 3 h, washed with Hanks' balanced salt solution (Life Technologies), fixed, and stained with 10% Giemsa prior to determination of adherence patterns by light microscopy.

Electron microscopy. Electron microscopy and immunogold electron microscopy were performed by standard methods with a JOEL JEM 1200 EX II transmission electron microscope (24). Antiserum to affinity-purified MBP-AggA fusion protein was used at a dilution of 1:100, and commercial antiserum to MBP (NEB) was used at the same dilution.

RESULTS

Tn*phoA* mutational analysis of AAF/I region 1. Results from Tn*phoA* mutagenesis of pJPN31 have previously been reported (29, 31). Of 86 nonadherent insertion mutants originally reported (29), 3 represented independent insertions into region 1. The precise insertion site was determined for each of these three mutants by restriction mapping and Tn*phoA* junctional DNA sequence analysis, and these sites are represented in Fig. 1B by square lollipops. One of these three mutants expressed PhoA activity (filled square in Fig. 1B).

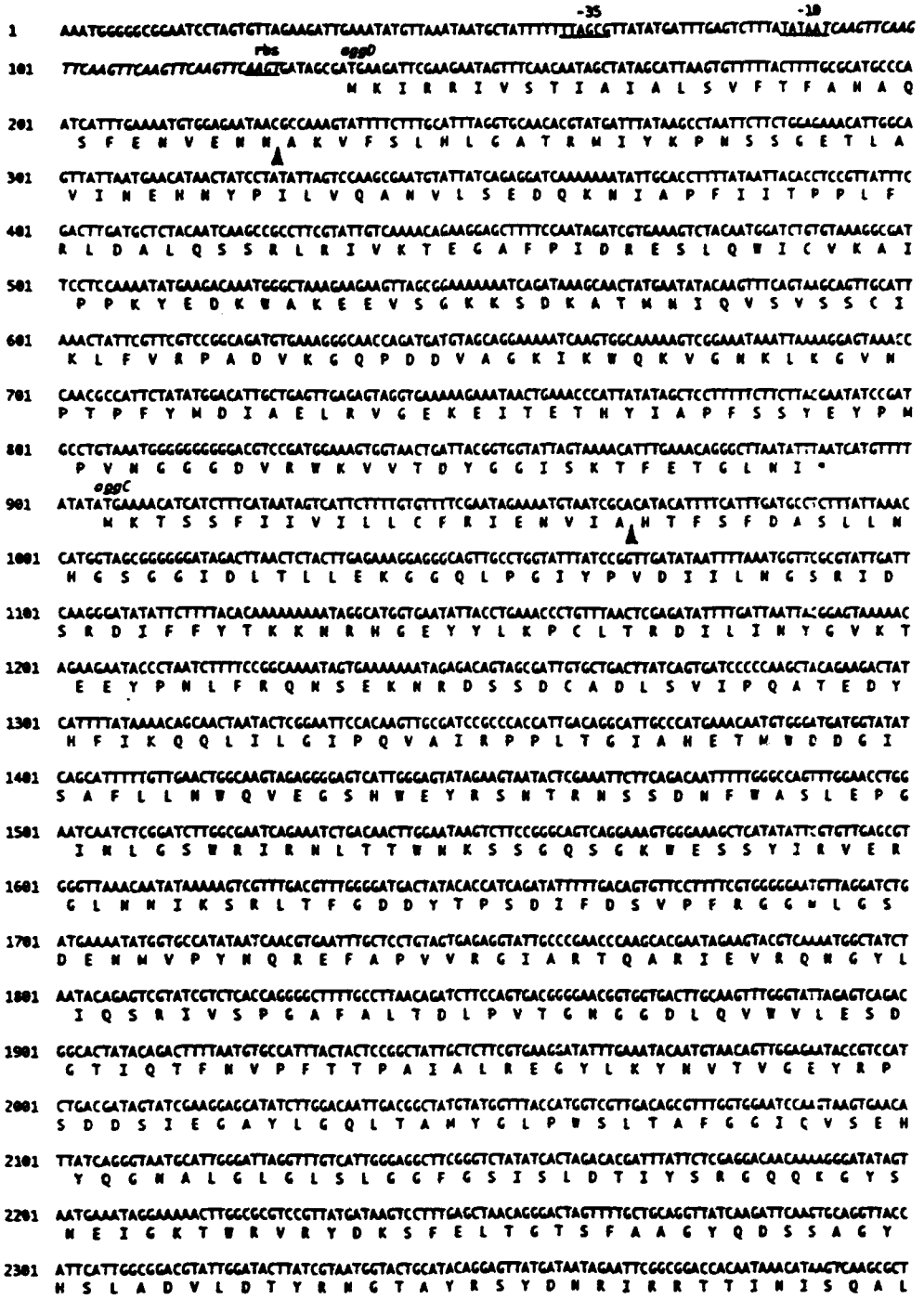
Tn*phoA* mutagenesis of HB101(pSS101) yielded six nonsibling PhoA$^+$ mutants (of 1,170 total PhoA$^+$ transconjugants) with single transposon insertions in pSS101. Three of these six mutants were unable to adhere to HEp-2 cells. Restriction mapping and Tn*phoA* junctional sequence analysis placed each of these three mutations in AAF/I region 1. The Tn*phoA* insertion sites for these three mutants are represented in Fig. 1B by filled triangular lollipops.

Tn*phoA* insertions into pJPN36 were screened for acquisition of Km^r without regard to PhoA activity. Of 40 such Km^r clones generated, 10 nonsibling mutants were found to be nonadherent to HEp-2 cells in an HB101(pJPN52) host. The exact location of Tn*phoA* insertion for each of these 10 mutants was determined as for the previously described mutants, and these sites are represented in Fig. 1 by circular lollipops.

In total, 16 Tn*phoA* mutational insertions were obtained in one of three constructions containing AAF/I region 1, each of which abrogated the AA pattern of adherence to HEp-2 cells shown by the parent construct. Within the 6.8-kb *Cla*I fragment originally reported as demarcating region 1 (31), all 16 mutational insertion sites are located within a 4.6-kb segment.

Nucleotide sequence of AAF/I region 1. The nucleotide sequence of both strands of the 6.8-kb *Cla*I fragment was determined. Translation of the DNA sequence showed four contiguous ORFs of greater than 100 codons, all within the 4.6-kb segment containing the AA$^-$ Tn*phoA* insertions. All four ORFs are oriented in the same direction, i.e., from left to right on the linear map shown in Fig. 1. These ORFs are given the genetic name *agg* (for aggregative adherence fimbria I), and the individual genes are designated (from left to right) *aggD*, *aggC*, *aggB*, and *aggA*. With the exception of *aggB*, Tn*phoA* insertions which abolished the AA phenotype were obtained in each of these ORFs, and the translated protein products show relatedness to components of other fimbrial biogenesis gene clusters. Figure 2 displays the DNA sequence of the 4.6-kb segment containing the *agg* gene cluster and the deduced amino acid sequence. The overall G+C content of the *agg* gene cluster is 43%.

***aggD* sequence.** Analysis of the nucleotide sequence encoding AggD revealed an ORF of 756 nucleotides which starts with the ATG codon at nucleotide 133 (Fig. 2). A stretch of



nucleotides consisting of a 6-bp segment directly repeated six times was detected in the noncoding region 5′ to *aggD*. A putative ribosome binding site and −10 and −35 sequences are also shown upstream of *aggD* in Fig. 2. Analysis of the encoded amino acid sequence by the algorithm of von Heijne (44) predicts the presence of an amino-terminal signal peptide of 32 amino acids. The predicted mature protein contains 220 amino acid residues and has a calculated molecular mass of 24.6 kDa and an isoelectric point of 8.9.

Alignment of the AggD amino acid sequence with proteins in the GenBank Entrez Protein database revealed significant homology with periplasmic chaperone proteins involved in the biosynthesis of several other fimbrial and nonfimbrial adhesins including *E. coli* NfaE (1) and PapD (25), *Yersinia pestis* Caf1M (12) and PsaB (26), and *Yersinia enterocolitica* MyfB (16). The percent amino acid identity between AggD and each of these reported chaperone proteins is shown in Table 1. An ideal alignment between AggD, NfaE, Caf1M, MyfB, and PapD is shown in Fig. 3. PapD, the prototype periplasmic pilus chaperone, binds P-pilus subunits and delivers them to the outer

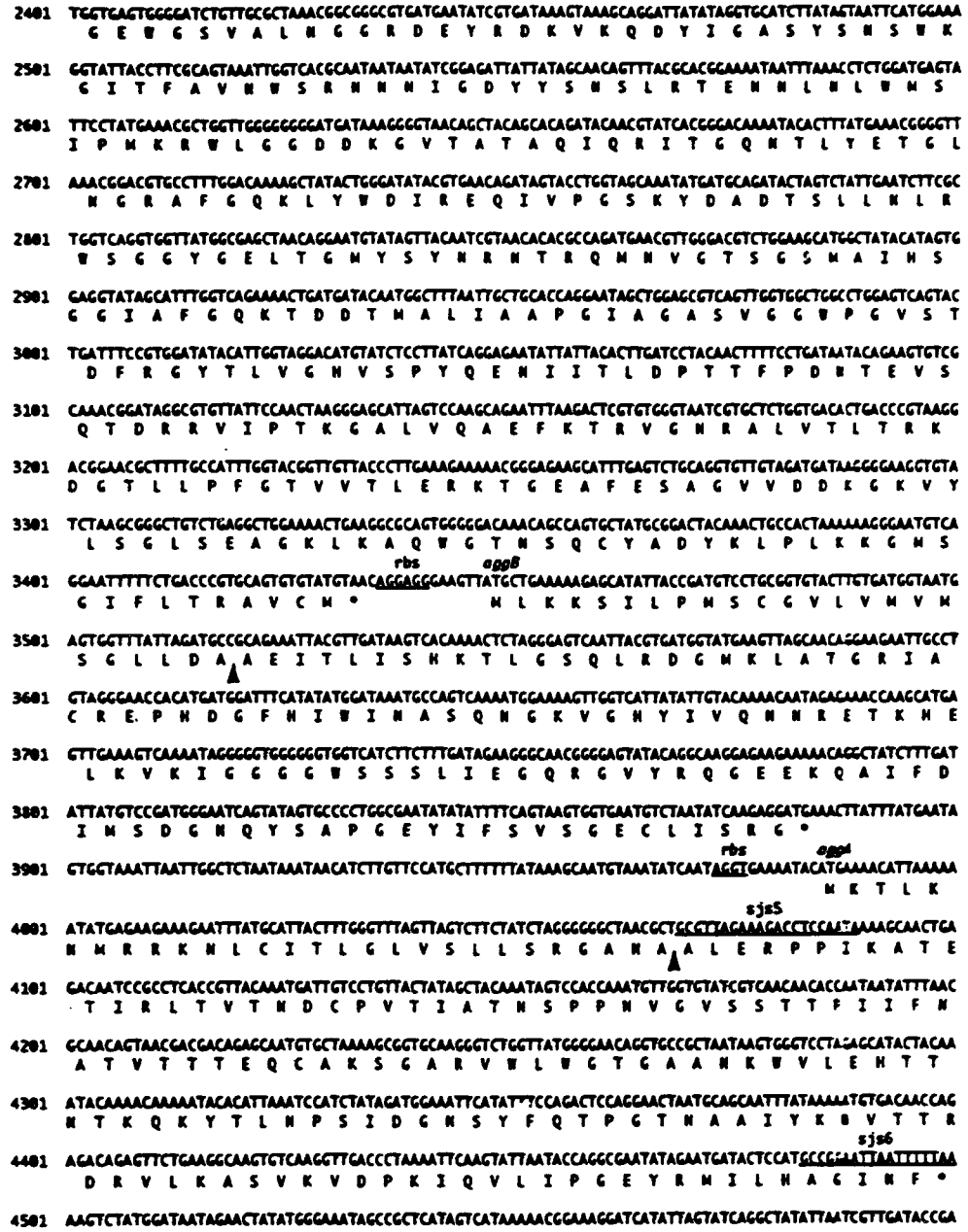


FIG. 2. Nucleotide and predicted amino acid sequences of the *agg* gene cluster (*aggDCBA*) within AAF/I region 1. The nucleotide coordinates are numerically indicated in the left margin. Single-letter codes for amino acids are shown below the corresponding nucleotide sequence (termination codons are shown as asterisks). Gene designations are noted above the start codons of the respective ORFs. Predicted −10 and −35 promoter sequences upstream of *aggD* and putative ribosome binding sites (rbs) are labelled and underlined. A stretch of 6 bp directly repeated six times in the noncoding region 5′ to *aggD* is italicized. The corresponding DNA sequences for oligonucleotide primers sjs5 and sjs6, used to amplify part of the *aggA* gene, are underlined and labelled. The predicted site of signal peptide cleavage for each protein, according to the formula of von Heijne (44), is marked by a vertical arrowhead.

membrane for assembly into pili (15). It has two globular domains (each consisting of multiple antiparallel β-strands) oriented towards one another to form a cleft and contains conserved and variable amino acids, features it has in common with the immunoglobulin superfamily (15). As noted in Fig. 3, AggD shows a high degree of similarity to the PapD consensus sequence, which includes amino acid residues important to the maintenance of proper tertiary structure and formation of the cleft wherein subunit binding occurs (15, 18). Distinct from the PapD prototype, AggD, NfaE, Caf1M, and MyfB have two cysteine residues, one each within the F1 and G1 β-strands (with reference to the solved structure of PapD). While cysteine residues are rare in this family of molecules, a pair of cysteine residues are found in the G2 and H2 β-strands of PapD and have been shown to link these two strands together (15).

TABLE 1. Identities between AggD or AggC and other pilus chaperone proteins or outer membrane usher proteins, respectively

Protein	Bacterial species	% identity to:		Amino acid overlap (no. of residues)
		AggC	AggD	
Chaperones				
NfaE	*E. coli*		65	202
Caf1M	*Y. pestis*		41	220
PsaB	*Y. pestis*		44	233
MyfB	*Y. enterocolitica*		41	220
PapD	*E. coli*		34	200
Usher proteins				
Caf1A	*Y. pestis*	42		811
MyfC	*Y. enterocolitica*	42		809
FimC	*B. pertussis*	27		811
PapC	*E. coli*	24		811

***aggC* sequence.** Seventeen base pairs downstream of the *aggD* ORF, an ORF of 2,526 nucleotides was detected and is designated *aggC*. No potential ribosome binding site was discerned upstream of the ATG initiation codon (at position 905). The *aggC* deduced gene product has a typical signal peptide of 21 amino acid residues. The expected mature protein has a calculated M_r of 90,900 and a pI of 7.4.

A protein homology comparison of AggC with the protein database revealed significant homology to outer membrane molecular usher proteins including *Y. enterocolitica* MyfC (16), *Y. pestis* Caf1A (19), and *Bordetella pertussis* FimC (45). A lower degree of homology is noted between AggC and PapC (33), the prototype outer membrane usher involved in P-pilus biogenesis. The percent amino acid identities between AggC and each of these usher proteins are shown in Table 1.

***aggB* sequence.** Seventeen base pairs downstream of the *aggC* stop codon, at position 3447 (Fig. 2), an ATG initiation codon for a third ORF is noted. This ORF has tentatively been designated *aggB*. A putative ribosome binding site is noted 7 bp upstream of the initiation codon. The predicted amino-terminal sequence of *aggB* contains a typical signal peptide of 24 amino acid residues, and the expected mature protein has a molecular mass of 13.3 kDa and a pI of 8.4. As noted above, no AA$^-$ Tn*phoA* insertion mutations were localized to the *aggB* ORF, and a homology search revealed no significant similarity to other reported proteins.

***aggA* sequence.** The *aggB* gene is followed by a 101-bp intergenic segment and then a 513-bp ORF, designated *aggA*. This ORF is preceded by a putative ribosome binding site 9 bp upstream of the ATG start codon at position 3985 (Fig. 2). The deduced AggA protein contains a predicted signal sequence of 28 amino acid residues, followed by a mature polypeptide of 143 residues with a molecular mass of 15.6 kDa and a theoretical isoelectric point of 9.7. The AggA protein exhibits nominal homology (17% amino acid identity) with the major subunit of F41 fimbriae (11) (data not shown).

Localization of the AggA gene product. As depicted in Fig. 1, the generation of three PhoA$^-$ Tn*phoA* insertions in frame with the *aggA* ORF confirmed that AggA is normally exported. Because AggA is required for expression of the AA phenotype and shows homology with the F41 major fimbrial subunit, it was hypothesized that this protein is a structural subunit of AAF/I fimbriae.

The ultrastructural location of AggA was determined by immunogold electron microscopy with antiserum raised against affinity-purified MBP-AggA fusion protein. This antiserum specifically decorated bundle-forming fimbriae (AAF/I) on wild-type EAggEC strain 17-2, as shown in Fig. 4. As a control, anti-MBP antiserum did not decorate the surface of strain 17-2. Additionally, MBP-AggA antiserum did not decorate 17-2(Δ*aggR*) (32), which does not express AAF/I (data not shown).

DISCUSSION

Nataro et al. (31) have previously described two noncontiguous plasmid regions required for AAF/I expression and AA. In this paper, we present a detailed analysis of the AAF/I determinants within region 1. These determinants consist of four contiguous genes within a 4.6-kb span of region 1 of the 60-MDa plasmid from wild-type EAggEC strain 17-2. On the basis of mutational analysis, amino acid sequence homologies, and ultrastructural data, it is inferred that the products of three of these four ORFs, namely, AggA, AggC, and AggD, serve as a structural fimbrial subunit, outer membrane usher protein, and periplasmic chaperone, respectively, in AAF/I biogenesis. The function of the *aggB* gene product could not be established from the studies presented herein. By using another strategy, however, it has since been shown that insertional mutations in this gene result in weak expression of the AA phenotype with HEp-2 cells yet abolish mannose-resistant hemagglutination (28a). This may suggest some differences in the genetic determinants required for the AA phenotype and mannose-resistant hemagglutination and is especially interesting since no proteins homologous to AggB have yet been reported for other fimbrial biogenesis systems.

In addition to the *agg* gene cluster described herein, a single determinant in AAF/I region 2 is essential for AAF/I production and AA. This gene, called *aggR*, is an *araC* homolog and appears to be a transcriptional activator of AAF/I expression (32). The plasmid pJPN52 contains the entire *aggR* gene and, as noted previously, complements in *trans* pJPN36, a plasmid construct containing AAF/I region 1, to allow expression of the AA phenotype. In this context, the directly repeated hexamer (TCAAGT) 5′ to the transcription start point of *aggD* may be important (Fig. 2). AraC induces transcription of P_{BAD} by binding to directly repeated half-sites (7). Likewise, other prokaryotic regulatory proteins, such as ToxR, also appear to recognize directly repeated DNA sequences (25). Studies are in progress to determine if this tandemly repeated hexamer is involved in AggR binding and transcriptional activation.

The low GC content of the *agg* gene cluster relative to that of the *E. coli* genome raises some questions as to its evolutionary origin. Of note, less than 1 kb upstream of the start codon of *aggD*, an ORF with a deduced amino acid sequence sharing 43% identity with the resolvase gene of Tn*3* (*tnpR*) was discerned, and a 38-bp segment flanking this ORF shares 58% identity with the Tn*3* inverted-repeat sequence (data not shown) (14). While there is a precedent for plasmid-encoded virulence determinants being found on transposons (40), the ability of the *agg* gene cluster to transpose has not been experimentally determined.

Organizationally, the *agg* gene cluster shares similarities with certain other fimbrial operons. Like CFA I of enterotoxigenic *E. coli* (46), the gene regulating transcription of fimbriae and accessory proteins is noncontiguous. Interestingly, the gene encoding EAggEC heat-stable enterotoxin 1 (EAST1) is located ≤1 kb downstream of the *aggA* stop codon, in the span of DNA between regions 1 and 2 (39), much like the organization of CFA/I determinants and the heat-stable enterotoxin (STa) gene on some plasmids in enterotoxigenic *E. coli* strains

(46). If direct data to support a role for AA and production of EAST1 in EAggEC disease pathogenesis are obtained, this span would appear to comprise an important virulence cassette.

The *aggA* gene, which appears to encode the major fimbrial subunit of AAF/I, is located downstream of other assembly genes. In this regard, it is organizationally similar to the operons involved in biogenesis of each of the adhesins in the Dr adhesin family (which recognize the Dr blood group antigen as a receptor), including F1845 fimbriae (5), Dr hemagglutinin (41), and the afimbrial adhesins AFA-I (22) and AFA-III (23). These are distinguished from operons encoding determinants of many other fimbriae, whereby the major fimbrial subunit gene is at or near the 5′ end of the operon (21). Likewise, the *agg* gene cluster is configured similarly to the determinants of members of the Dr adhesin family with respect to the overall number and relative order of genes required for adhesin expression (5, 22, 23, 41). Distinct from the AAF/I determinants, though, the presumed regulatory genes of the Dr adhesin family are contiguous with the genes encoding the assembly proteins and adhesin subunit, usually at the 5′ end of the operon (5, 22).

Three distinct mannose-resistant patterns of adherence to HEp-2 cells have been discerned in studying so-called enteroadherent *E. coli*: LA, AA, and DA (38). LA is characteristic of EPEC. Evidence to date suggests that bundle-forming pilin, encoded by *bfpA*, may represent the LA adhesin (10). Ninety-nine percent of EPEC strains hybridized with a *bfpA* gene probe, suggesting a common genetic determinant mediating this phenotype (13). BFP pilin shares characteristic features with other pilins such as *Vibrio cholerae* TcpA and *Neisseria gonorrheae* MS11, all members of the type IV pilin family (10).

While two different DA determinants have been cloned, the F1845 fimbrial adhesin has been most thoroughly characterized (4, 5). DNA hybridization studies show that approximately 60 to 80% of DAEC strains share relatedness with F1845 genetic determinants (28a). As mentioned above, F1845 belongs to the family of Dr adhesins. Like other members of this family, F1845 is both the major fimbrial subunit and the DA adhesin (34).

From the data presented herein, it is clear that the AAF/I determinants are genetically distinct from those determinants described for the LA and DA phenotypes. On the basis of

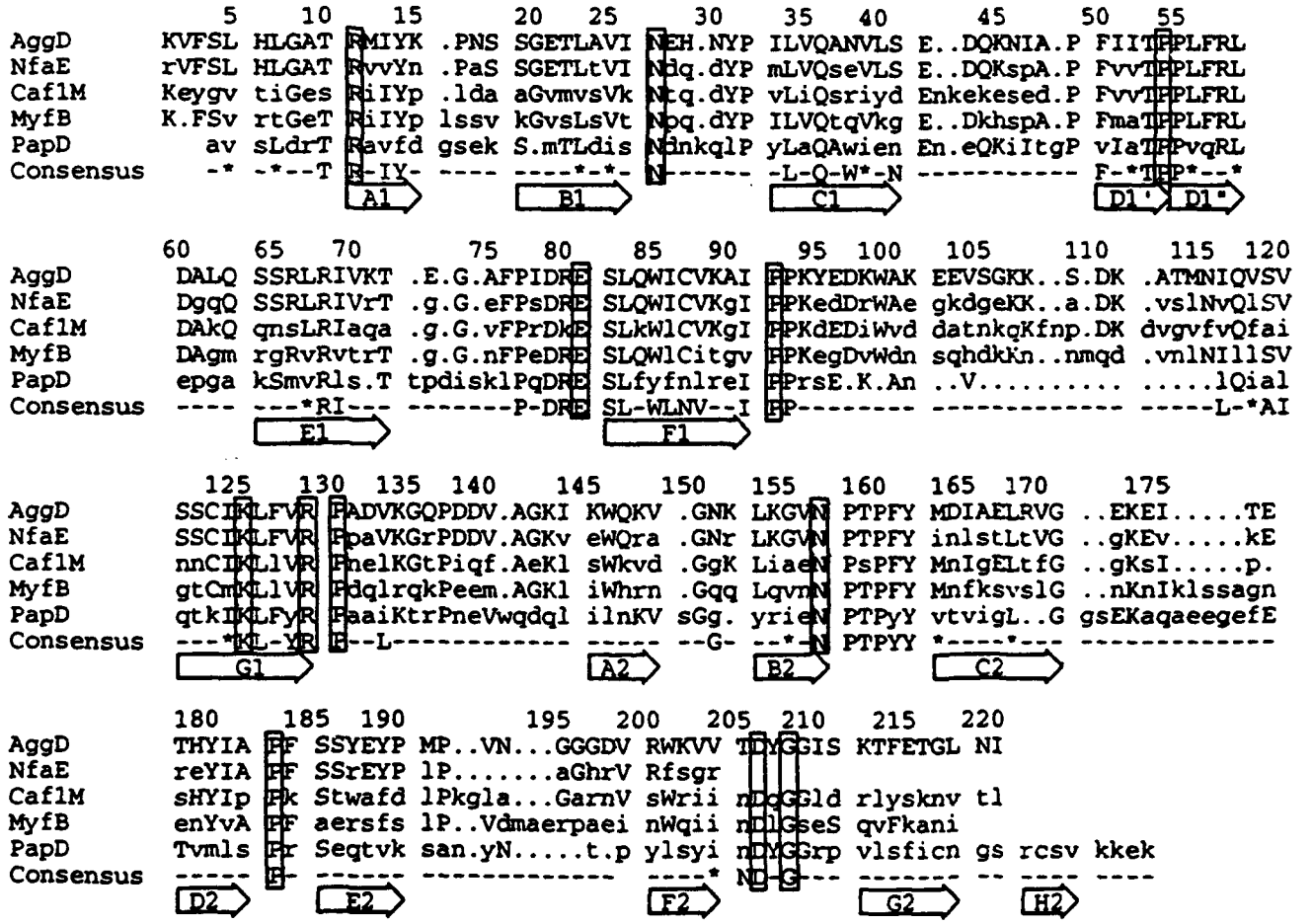


FIG. 3. The inferred primary sequence of AggD is shown in alignment with the primary sequences of NfaE (11), Caf1M (12), MyfB (16), PapD (25), and the chaperone consensus sequence (17). Residues in the alignment are numbered according to AggD. Capital letters indicate amino acids identical to the corresponding residues in AggD. The consensus sequence was previously compiled from a comparison of 13 members of the chaperone family (17) not including AggD, NfaE, and MyfB. Asterisks in the consensus indicate positions of conserved hydrophobic character. Capital letters in the consensus indicate residues conserved in 8 of 13 sequences, while boxed residues are invariant for all 16 chaperone proteins. Periods represent gaps introduced for optimal alignment. The labelled arrows below the sequences indicate the β-strands resolved from the crystal structure of PapD, which are arranged in sheets to form two globular domains.

FIG. 4. Immunogold electron micrograph of strain 17-2 decorated with gold particles after primary incubation with anti-MBP-AggA antiserum. Bar, 200 nm.

nucleotide and deduced amino acid sequences of the AggA fimbrial subunit. AAF I does not readily fall into any known family of fimbriae, although modest homology with F41 is recognized (11). Furthermore, while it appears that AggA represents a major AAF/I fimbrial subunit, data to assign AggA a role as the actual AA adhesin are as yet lacking. It is appreciated that the AA phenotype generally presents three different interactions. Bacteria adhere not only to the epithelial cell but also to one another (interbacterial interaction) and often to the glass matrix. Although the *agg* gene cluster is required for expression of the AA phenotype, the molecular basis for each of these interactions is not yet clear. We have produced antiserum against AggA which specifically decorates AAF/I fimbriae in vitro. Preliminary experiments suggest that this antiserum may inhibit HEp-2 cell adherence, although more detailed studies are ongoing to determine the exact role of AggA in HEp-2 cell AA.

ACKNOWLEDGMENTS

This work was supported by U.S. Navy Research and Development Research Work Unit 611020.B998.AKS131.A.1395 (S.J.S.) and by grant 1R29-AI-33907 through the National Institutes of Health (J.P.N.).

We thank Patricia Guerry and James B. Kaper for extensive advice and assistance and Daniel Scott for critical reading of the manuscript.

REFERENCES

1. **Ahrens, R., M. Ott, A. Ritter, H. Hoschutzky, T. Buhler, F. Lottspeich, G. J. Boulnois, K. Jann, and J. Hacker.** 1993. Genetic analysis of the gene cluster encoding nonfimbrial adhesin I from an *Escherichia coli* uropathogen. Infect. Immun. 61:2505–2512.
2. **Applied Biosystems, Inc.** 1992. High-quality template DNA for *Taq* cycle sequencing using DyeDeoxy terminators: an improved preparation procedure. Use: Bulletin 18:1–4.
3. **Ausubel, R. M., R. Brent, R. E. Kingston, D. D. Moore, J. G. Seidman, J. A. Smith, and K. Struhl.** 1989. Current protocols in molecular biology. John Wiley & Sons, New York.
4. **Benz, I., and M. A. Schmidt.** 1992. Isolation and serologic characterization of AIDA-I, the adhesin mediating the diffuse adherence phenotype of the diarrhea-associated *Escherichia coli* strain 2787 (O126:H27). Infect. Immun. 60:13–18.
5. **Bilge, S. S., C. R. Clausen, W. Lau, and S. L. Moseley.** 1989. Molecular characterization of a fimbrial adhesin, F1845, mediating diffuse adherence of diarrhea-associated *Escherichia coli* to HEp-2 cells. J. Bacteriol. 171:4281–4289.
6. **Birnboim, H. C., and J. Doly.** 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. Nucleic Acids Res. 7:1515–1523.
7. **Carra, J. H., and R. F. Schleif.** 1993. Variation of half-site organization and DNA looping by AraC protein. EMBO J. 12:35–44.
8. **Cravioto, A., R. J. Gross, S. Scotland, and B. Rowe.** 1979. An adhesive factor found in strains of *Escherichia coli* belonging to the traditional enteropathogenic serotypes. Curr. Microbiol. 3:95–99.
9. **Cravioto, A., A. Tello, A. Navarro, J. Ruiz, H. Villafan, F. Uribe, and C. Eslava.** 1991. Association of *Escherichia coli* HEp-2 cell adherence patterns with type and duration of diarrhoea. Lancet 337:262–264.
10. **Donnenberg, M. S., J. A. Giron, J. P. Nataro, and J. B. Kaper.**

1991. A plasmid-encoded type IV fimbrial gene of enteropathogenic *Escherichia coli* associated with localized adherence. Mol. Microbiol. 6:3427–3437.
11. **Fidock, D. A., P. A. McNicholas, and P. R. Lehrbach.** 1989. Nucleotide sequence of the F41 fimbriae subunit gene in *Escherichia coli* B41. Nucleic Acids Res. 17:2849.
12. **Galyov, E. E., A. V. Karlishev, T. V. Chernovskaya, D. A. Dolgikh, O. Y. Smirnov, K. I. Volkovoy, V. M. Abramov, and V. P. Zav'yalov.** 1991. Expression of the envelope antigen F1 of *Yersinia pestis* is mediated by the product of Caf1M gene having homology with the chaperone protein PapD of *Escherichia coli*. FEBS Lett. 286:79–82.
13. **Giron, J. A., M. S. Donnenberg, W. C. Martin, K. G. Jarvis, and J. B. Kaper.** 1993. Distribution of the bundle-forming pilus structural gene (*bfpA*) among enteropathogenic *Escherichia coli*. J. Infect. Dis. 168:1037–1041.
14. **Heffron, F., B. J. McCarthy, H. Ohtsubo, and E. Ohtsubo.** 1979. DNA sequence analysis of the transposon Tn3: Three genes and three sites involved in transposition of Tn3. Cell 18:1153–1163.
15. **Holmgren, A., M. J. Kuehn, C.-I. Branden, and S. J. Hultgren.** 1992. Conserved immunoglobulin-like features in a family of periplasmic pilus chaperones in bacteria. EMBO J. 11:1617–1622.
16. **Iriarte, M., J. Vanooteghem, I. Delor, R. Diaz, S. Knutton, and G. R. Cornelis.** 1993. The Myf fibrillae of *Yersinia enterocolitica*. Mol. Microbiol. 9:507–520.
17. **Jacob-Dubuisson, F., M. J. Kuehn, and S. J. Hultgren.** 1993. A novel secretion apparatus for the assembly of adhesive bacterial pili. Trends Microbiol. 1:50–55.
18. **Jones, C. H., J. S. Pinkner, A. V. Nicholes, L. N. Slonim, S. N. Abraham, and S. J. Hultgren.** 1993. FimC is a periplasmic PapD-like chaperone that directs assembly of type 1 pili in bacteria. Proc. Natl. Acad. Sci. USA 90:8397–8401.
19. **Karlyshev, A. V., E. E. Galyov, O. Y. Smirnov, A. M. Guzayev, V. M. Abramov, and V. P. Zav'yalov.** 1992. A new gene of the f1 operon of *Y. pestis* involved in the capsule biogenesis. FEBS Lett. 297:77–80.
20. **Knutton, S., R. K. Shaw, M. K. Bhan, H. R. Smith, M. McConnell, T. Cheasty, P. H. Williams, and T. J. Baldwin.** 1992. Ability of enteroaggregative *Escherichia coli* strains to adhere in vitro to human intestinal mucosa. Infect. Immun. 60:2083–2091.
21. **Krogfelt, K. A.** 1991. Bacterial adhesion: genetics, biogenesis and role in pathogenesis of fimbrial adhesins of *Escherichia coli*. Rev. Infect. Dis. 13:721–735.
22. **Labigne-Roussel, A. F., M. A. Schmidt, W. Walz, and S. Falkow.** 1985. Genetic organization of the *afa* operon and nucleotide sequence from a uropathogenic *Escherichia coli* gene encoding an afimbrial adhesin (AFA-1). J. Bacteriol. 162:1285–1292.
23. **Le Bouguenec, C., M. I. Garcia, V. Ouin, J. M. Desperrier, P. Gounon, and A. Labigne.** 1993. Characterization of plasmid-borne *afa-3* gene clusters encoding afimbrial adhesins expressed by *Escherichia coli* strains associated with intestinal or urinary tract infections. Infect. Immun. 61:5106–5114.
24. **Levine, M. M., P. Ristaino, G. Marley, C. Smyth, S. Knutton, E. Boedeker, R. Black, C. Young, M. L. Clements, C. Cheney, and R. Patnaik.** 1984. Coli surface antigens 1 and 3 of colonization factor antigen II-positive enterotoxigenic *Escherichia coli*: morphology, purification, and immune responses in humans. Infect. Immun. 44:409–420.
25. **Lindberg, F., J. M. Tennent, S. J. Hultgren, B. Lund, and S. Normark.** 1989. PapD, a periplasmic transport protein in P-pilus biogenesis. J. Bacteriol. 171:6052–6058.
26. **Lindler, L. E., and B. D. Tall.** 1993. *Yersinia pestis* pH6 antigen forms fimbriae and is induced by intracellular association with macrophages. Mol. Microbiol. 8:311–324.
27. **Mandel, M., and A. Higa.** 1970. Calcium dependent bacteriophage DNA infection. J. Mol. Biol. 53:159–162.
28. **Miller, V. L., R. K. Taylor, and J. J. Mekalanos.** 1987. Cholera toxin transcriptional activator *toxR* is a transmembrane DNA binding protein. Cell 48:271–279.
28a. **Nataro, J. P.** Unpublished data.
29. **Nataro, J. P., Y. Deng, D. R. Maneval, A. L. German, W. C. Martin, and M. M. Levine.** 1992. Aggregative adherence fimbriae I of enteroaggregative *Escherichia coli* mediate adherence to HEp-2 cells and hemagglutination of human erythrocytes. Infect. Immun. 60:2297–2304.
30. **Nataro, J. P., J. B. Kaper, R. Robins-Browne, V. Prado, and M. M. Levine.** 1987. Patterns of adherence of diarrheagenic *Escherichia coli* to HEp-2 cells. Pediatr. Infect. Dis. J. 6:829–831.
31. **Nataro, J. P., D. Yikang, J. A. Giron, S. J. Savarino, M. H. Kothary, and R. Hall.** 1993. Aggregative adherence fimbriae I expression in enteroaggregative *Escherichia coli* requires two unlinked plasmid regions. Infect. Immun. 61:1126–1131.
32. **Nataro, J. P., D. Yikang, D. Yinkang, and K. Walker.** 1994. AggR, a transcriptional activator of aggregative adherence fimbria I expression in enteroaggregative *Escherichia coli*. J. Bacteriol. 176:4691–4699.
33. **Norgren, M., M. Baga, J. M. Tennent, and S. Normark.** 1987. Nucleotide sequence, regulation and functional analysis of the *papC* gene required for cell surface localization of Pap pili of uropathogenic *Escherichia coli*. Mol. Microbiol. 1:169–178.
34. **Nowicki, B., A. Labigne-Roussel, S. Moseley, R. Hull, S. Hull, and J. Moulds.** 1990. The Dr hemagglutinin, AFA-I, AFA-III, and F1845 fimbriae of uropathogenic *Escherichia coli* belong to a family of hemagglutinins with Dr receptor recognition. Infect. Immun. 58:279–281.
35. **Robinson, M. K., P. M. Bennett, S. Falkow, and H. M. Dodd.** 1980. Isolation of a temperature sensitive derivative of RP1. Plasmid 3:343–347.
36. **Sambrook, J., E. F. Fritsch, and T. Maniatis.** 1989. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y.
37. **Sanger, F., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA 74:5463–5467.
38. **Savarino, S. J.** 1993. Enteroadherent *Escherichia coli*: a heterogeneous group of *E. coli* associated with diarrhoeal diseases. 1993. Trans. R. Soc. Trop. Med. Hyg. 87 Suppl. 3):49–53.
39. **Savarino, S. J., A. Fasano, J. Watson, B. M. Martin, M. M. Levine, S. Guandalini, and P. Guerry.** 1993. Enteroaggregative *Escherichia coli* heat-stable enterotoxin 1 represents another subfamily of *E. coli* heat-stable toxin. Proc. Natl. Acad. Sci. USA 90:3093–3097.
40. **So, M., and B. J. McCarthy.** 1980. Nucleotide sequence of the bacterial transposon Tn1681 encoding a heat-stable toxin and its identification in enterotoxigenic *Escherichia coli* strains. Proc. Natl. Acad. Sci. USA 77:4011–4015.
41. **Swanson, T. N., S. S. Bilge, B. Nowicki, and S. L. Moseley.** 1991. Molecular structure of the Dr adhesin: nucleotide sequence and mapping of receptor-binding domain by use of fusion constructs. Infect. Immun. 59:261–268.
42. **Taylor, R. K., C. Manoil, and J. J. Mekalanos.** 1989. Broad-host-range vectors for delivery of TnphoA: use in genetic analysis of secreted virulence determinants of *Vibrio cholerae*. J. Bacteriol. 171:1870–1878.
43. **Vial, P. A., R. Robins-Browne, H. Lior, V. Prado, J. B. Kaper, J. P. Nataro, D. Maneval, A. Elsayed, and M. M. Levine.** 1988. Characterization of enteroadherent-aggregative *Escherichia coli*, a putative agent of diarrheal disease. J. Infect. Dis. 158:70–79.
44. **Von Heijne, G.** 1983. Patterns of amino acids near signal sequence cleavage sites. Eur. J. Biochem. 133:17–21.
45. **Willems, R. J. L., H. G. J. van der Heide, and F. R. Mooi.** 1992. Characterization of a *Bordetella pertussis* fimbrial gene cluster which is located directly downstream of the filamentous haemagglutinin gene. Mol. Microbiol. 6:2661–2671.
46. **Willshaw, G. A., H. R. Smith, and B. Rowe.** 1985. Cloning of regions encoding colonisation factor antigen 1 and heat-stable enterotoxin in *Escherichia coli*. FEMS Microbiol. Lett. 16:101–106.
47. **Yamamoto, T., S. Endo, T. Yokota, and P. Echeverria.** 1991. Characteristics of adherence of enteroaggregative *Escherichia coli* to human and animal mucosa. Infect. Immun. 59:3722–3737.
48. **Yamamoto, T., Y. Koyama, M. Matsumoto, E. Sonoda, S. Nakayama, M. Uchimura, W. Paveenkittiporn, K. Tamura, T. Yokota, and P. Echeverria.** 1992. Localized, aggregative, and diffuse adherence to HeLa cells, plastic, and human small intestines by *Escherichia coli* isolated from patients with diarrhea. J. Infect. Dis. 166:1295–1310.

A-1 | 20